



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

언론정보학과 석사학위 논문

Interaction Behavior as an Indicator of “Online Homophily”:

A Study on Intimacy among Facebook Friendship

(상호작용 행위를 통해 본 “온라인 호모필리”

: 페이스북 “친구” 관계의 친밀도에 대한 연구)

서울대학교 대학원

언론정보학과

위지은

Abstract

Homophily refers to the phenomenon of people associating more with similar people than they do with non-similar or dissimilar people. While most of the research on homophily has been focused on offline contexts, it is necessary to study the phenomenon in online contexts as well since people are "always on." Exploring homophily in Facebook is particularly important in that, though Facebook users primarily tend to maintain or solidify already-existing offline relationships through the online space, Facebook "friendships" indicate a broader spectrum of relationships ranging from complete strangers to confiding relations. Nevertheless, observation on homophily in computer-mediated contexts has been conducted through measuring the effect in the same way as in offline contexts. Thus, current research aims to test whether the traditional concept of homophily is applicable to "online homophily" with further speculation by collecting users' profile information and activity log data, in addition to users' survey response data. Results show that empirical evidence on homophily in the online space is not found based on the original concept of similarities on personal attributes, but through interaction behaviors between users. Even though the significance of correlations between interaction behaviors and perceived intimacy was not so strong,

findings from current research imply that further elaboration on “online homophily” is necessary, in a way more relevant to the computer-mediated context other than the traditional concept of homophily.

Key words: homophily, computer-mediated communication (CMC), intimacy, social network services (SNSs), Facebook, user profile, similarity, interaction behavior

Student Number: 2011-23156

Table of Contents

I. INTRODUCTION	1
II. LITERATURE REVIEW	6
1. Homophily	6
2. Facebook	11
1) Facebook User Information	14
2) Interaction Behaviors in Facebook.....	15
3. Limitations on the Literature	18
III. RESEARCH METHOD.....	22
1. Issues on Facebook User Data Collection	23
2. Participant Recruitment	26
3. Survey.....	26
4. Experiment Procedure	28
5. Data Crawling	31
1) Similarity.....	31
2) Interaction.....	33
IV. RESULTS	36
1. Descriptive Analysis	36
2. Multiple Linear Regression (MLR)	37
1) Similarity on Attributes that Composes Intimacy	38
2) Interaction Types that Predict Intimacy	40
V. DISCUSSION.....	46
VI. APPENDIX	57
VII. REFERENCE	60

List of Tables and Figures

Table 1. Similarity Variables

Table 2. Interactions on Timeline

Table 3. Interaction on Messenger (Chat/ Message)

Table 4. Basic Information of Participants

Table 5. MLR Analysis Result: Similarity

Table 6. Descriptive Analysis on Interaction of 36 Participants

Table 7. MLR Analysis Result: Interaction (Linear-Log Model)

Figure 1. Example on the Survey Questionnaire

Figure 2. Request on access to participant's Facebook data

**Figure 3. Plot on Intimacy and Logarithmic Transformation on
Interaction (Simple Count) of 971 Relationship**

I. INTRODUCTION

People form and maintain relationships with one another. Yet, not all the relationships are identical; they differ in the degree of intimacy and intensity. Among many possible explanations on why people show different communication behaviors in different relationships, an extensive literature established by social scientists manifests an approach towards the fundamental question on humanity, referring to as ‘homophily.’ Probably the simplest way to describe homophily is by using the proverbial expression, “birds of a feather flock together.” Homophily principle is the tendency for people sharing similar traits to affiliate more with each other, than with people sharing dissimilar traits or not sharing similar traits. Specifically, it means that people tend to feel more attached to others with similar attributes, associate more often with the similar people, and as a result, develop closer relationships with them. From a perspective of empirically looking into relationship as a dynamic process rather than a static state, keeping relations could be observed in terms of seeking continued contact with the counterpart (Lazarsfeld, & Merton, 1954).

A large volume of research on elaborating the concept and principle of ‘homophily’ has been conducted within the offline context,

which is natural considering the fact that the Internet became popular only after the advent of World Wide Web in the 1990s. Studies of homophily within online contexts were often focused on particular situations, such as online dating, looking into characteristics related to certain purposes (Fiore, & Donath, 2005). In such case, personal attributes that are examined are very specific, as “marital status,” “wants children,” “pet preferences,” which could provide articulated explanations on homophily among marriage relationships. However, those attributes alone are difficult to explain homophily effect among people in general. Other research examined homophily in a more general condition through MySpace, investigating the homophily effect of respective attributes one by one such as sex homophily, ethnic homophily, not as a whole (Thelwall, 2009).

It is crucial to understand communication behaviors of people happening within the online context in general because people are “always on” (Baron, 2008). They exchange information, debate controversial issues, and have casual conversations online as they do face-to-face. Social Network Services (SNSs) such as Twitter and Facebook facilitate such interactions. While relationships built in Twitter as “following” and being followed do not require technical reciprocation and often are not directly connected to offline relationship, relationships on Facebook require mutual confirmation on

relationships and are often based on offline relationships (Marwick, & boyd, 2010). Most of the Facebook users tend to maintain or solidify existing offline relationships through the service.

Due to the feature, elaboration on homophily in Facebook is particularly important because the observation may provide insights on the similarities and differences in communication behaviors between online and offline contexts. Not only does Facebook reflect users' offline relationships within the computer-mediated context, but also show distinct online communication behaviors such as "messaging," "liking" and "sharing." Moreover, Facebook has more than one billion users, and thus, stores enormous information on relationship formation and maintenance behaviors among the users.¹ Thus careful examination on homophily phenomenon on Facebook seems crucial (Hampton et al., 2011; Lampe et al., 2006, 2007; Sheldon, 2008; Wilson et al., 2012).

In order to study homophily within the context of online, the scope of understanding the phenomenon has to be expanded. Unlike in-person relationships, online relationships are often formed between people who have never met each other in person before. For example, Facebook 'friends' are different from friends that we ordinarily refer to

¹ One Billion Fact Sheet, Facebook Newsroom. (October 4, 2012) Available at: <http://newsroom.fb.com/imagelibrary/downloadmedia.ashx?MediaDetailsID=4227&SizeId=-1>

in our everyday lives. Though personal networks on Facebook are based on already-existing relationships, Facebook ‘friends’ indicate people connected through the service including complete strangers, latent ties, and close friends, traditional meaning of friends implicates mutual affection addition to such connection. In other words, it could be understood that ‘friends’ formed on Facebook cover a broader spectrum of relationship than friends within offline contexts (Ellison et al., 2010; Hampton et al., 2011; Marwick, & boyd, 2010). Since people do form relationships based on different levels of intimacy or sometimes even without any intimacy, yet are all still called equally as ‘friends’ on Facebook, looking beyond the explicit ‘friends’ lists on Facebook is necessary in order to understand relationships on Facebook when studying homophily.

As Lazarsfeld and Merton (1954) emphasized, relationships should be understood as a dynamic process that requires continued contact. Fortunately, interaction data between Facebook users and their ‘friends’ makes the empirical observation on homophily phenomenon as a dynamic process feasible. By looking into both the similarities and interaction behaviors of people on Facebook, whether the traditional concept of homophily is applicable to the online context is explored, and further, how diverse interaction behaviors relate to the phenomenon is investigated in the research.

Therefore, because the way people affiliate online differs from the way they do offline, we assume that the “online homophily” should be contemplated and measured distinctively from the homophily phenomenon, which has been studied mostly within the offline context. In other words, the perspective on homophily phenomenon simply as the perceived level of intimacy as an indicator of “closer relationships” deserves to be reconsidered in the computer-mediated context. Current study primarily focuses on providing evidence of homophily on Facebook, by collecting tremendous data of Facebook users including individual profiles and diverse activity behaviors.

First of all, empirical evidence on the original concept of homophily in the computer-mediated context is probed (Research Question 1), and secondly, the correlation between self-reported survey data on intimacy and the actual behavioral data as interactions on Facebook are explored (Research Question 2). The second approach is particularly meaningful in that, by investigating the relationship between a perceived intimacy and various interaction behaviors online, it made an attempt to provide an alternative or a complementary way to measure homophily in the computer-mediated context.

II. LITERATURE REVIEW

1. Homophily

Though there has been a number of studies related to the phenomenon, the term ‘homophily’ was first adopted to the academics by Lazarsfeld and Merton (1954) in their study of friendship formation processes. Research such as similarity-attraction hypotheses—claiming that people are generally attracted to those similar to themselves—, self-categorization theories—contesting that people appraise others as similar and dissimilar to themselves by categorizing their own attributes and comparing such with others’ attributes—, and studies probing similarity as a variable on the social tie formation are all in the context of homophily as well (Byrne, 1971; Homans, 1950; recited from Monge & Contractor, 2003; Verbrugge, 1977).

Lazarsfeld and Merton (1954) conceptualized the term homophily as a tendency in order to encourage the idea that different degrees of homophily among different relationships exist, and that the simple expression of “flocking together” lacks clarity in the terminology; thus, that positive correlations between the attribute similarity of people and relationships could be measured. Accordingly, the concept presumed that the effect of homophily would be greater when people share multiple attributes concurrently than when they share a single attribute.

Need for exploration and further testing on such comprehensive measure of homophily was manifested in the study.

The earliest studies of homophily focused on small social groups of college students, adolescents from school classes, and urban neighborhoods through ethnographic observations in the early 1900s (McPherson et al., 2001). Throughout the mid 1900s, though subjects of homophily research expanded to other public places, it still had restrictions on the observable range since the studies depended solely on observation of researchers. Modern sample surveys were applied to the studies in the 1970s. This methodological technology not only made empirical research on homophily among the large-scale groups feasible, e.g. schools, communities, but also made simultaneous measure on multiple characteristics of individuals possible; in a self-reporting way (Fischer, 1982; Shrum et al., 1988). Since the late 1980s, many studies have put efforts to provide more concrete empirical evidence of homophily through longitudinal research. Hallinan and Smith (1985) studied how inter-racial friendships among classmates change by self-reported friendship nomination of participants. Further, Burt (2000) conducted research on how relationships among colleagues decay throughout time by self-reporting surveys. Recent works in homophily made attempts to collect and analyze computer log data by looking directly into

communication behaviors, such as how often people associate with whom in which ways (Choudhury et al., 2010; Kossinets & Watts, 2009; Thelwall, 2009). The efforts enabled researchers to minimize biases that may occur from ethnographic observations or self-reporting sample survey methods as well.

Among diverse attributes explored through previous research as an evidence of homophily, salient dimensions are identified in the following:

Race / Ethnicity Race and ethnicity are the most salient dimensions that drive homophily across a wide range of relationships, from the most intimate relationships of marriage to merely acquaintances (Marsden, 1988; recited from McPherson et al., 2001; Shrum et al., 1988). Also in the case of computer-mediated circumstance such as MySpace, strong homophily was derived from ethnicity (Thelwall, 2009).

Sex / Gender Unless in sex-segregated structures, homophily effect created by sex and gender is considered far less than by race and ethnicity (McPherson & Smith-Lovin, 1987; McPherson et al., 2001). In recent studies, while formation of university students' relationships through e-mail showed positive sex homophily, user interactions through commenting in MySpace showed no sex homophily effect (Thelwall, 2009; Kossinets & Watts, 2009).

Age Age homophily is mostly found high in close friendships, superficial relationships such as mere acquaintances, and even in confiding relations, when not involving kin (Fischer, 1982; Marsden, 1988). Because schools group students together in similar ages together, in their school years, children show powerful homophily of age. However, such homophily weakens, as people get older (Shrum et al., 1988). Overall, age has a strong homophily effect, especially up to the age of 40 (Kossinets & Watts, 2009; Marsden, 1988; McPherson et al., 2011; Thelwall, 2009; recited from Yoon, 2006).

Religion Though not as strong as race and ethnicity, religion also has a significant effect on homophily (Fischer, 1982; Marsden 1988; Thelwall, 2009). The degree differed among religions: Catholics showed higher tendency than the Protestants. Also, the less close or intimate the relations, the lower the tendency of religious homophily (recited from McPherson et al., 2001).

Education / Occupation / Social Class Highly due to the strong influence of propinquity set by the surroundings, education, occupation, and social class generate great homophily; education on school setting, occupation on workplace, and social class on residential area (McPherson et al., 2001). Education inbreeds great level of homophily, though less than race and ethnicity (Marsden, 1988; recited from Yoon, 2006). In the study of Kossinets and Watts (2009),

specifically defined as undergraduate, graduate, non-degree, professional students, faculty, administrator, and staff within a university, occupation also appeared to have homophily effect. Some studies show that education, occupation, and social class tend to have greater effect on less intimate relationships (recited from McPherson et al., 2001; Verbrugge, 1977).

Values Values include attitudes, abilities, beliefs, and aspirations of people that determine their behaviors in the future. Favorable attitude, behavioral involvement, and perceived mutual identity are considered to lead to attraction and interaction (Huston & Levinger, 1978). Though literature demonstrates influence of political views on homophily, it is unclear whether such association is a direct consequence of political similarity or a consequence of other attributive similarities (Park, 2007; Verbrugge, 1977). Nevertheless, overall value-homophily tend to prevail in relationships regardless of relationship stages, because without sharing certain value such as political or racial attitudes, relationships will eventually dissolve (Lazarsfeld & Merton, 1954).

Based on the literature of prominent dimensions on homophily, which were based on observations offline, there has been research on homophily in the online context recently. Thelwall (2009) conducted an empirical research by collecting user data from MySpace in order to

look into the impact of personal attributes such as ethnicity, religion, age, and sex. Results demonstrate the prosperity of traditional sources of homophily in the online world; overall positive effect on homophily except for sex. In the study, Thelwall observed homophily effect through direct interaction behaviors, i.e. commenting on others' profile pages, unlike the original concept of homophily; which was measured by perceived intimacy and affiliation towards others, not the interaction behaviors. Also, Kossinets and Watts (2009) measured homophily by interaction, i.e. e-mail exchanges, in order to identify whether the association of people originated from the homophily effect—which assumes individual preference to define relationships—or from the structural foci surrounding individuals. In the study of Choudhury et al. (2010), homophily was defined as the similarity itself (not as the tendency of affiliation among the similar). With collected data from Twitter, they observed how attribute similarities between people influence information diffusion, according to information themes, e.g. politics, entertainment, technology.

2. Facebook

While the nature and specifics of social network services/sites (SNSs) may not be the same, e.g. Facebook, Twitter, MySpace, they

share a common ground. boyd and Ellison (2007) provided a definition on SNSs that explains such commonality as:

Web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by other within the system.

Specific functions of each service such as “poking” in case of Facebook, and decorating individual profiles in MySpace are features that make each service distinct. Among such various SNSs, Facebook became the number one around the world. Since its launch in 2004, Facebook has expanded in terms of its users as well as its influence to the society. By 2010, Facebook became the most visited website in the United States, and by 2012, the number of Facebook users surpassed one billion, which is a greater number than the population of Europe (recited from Wilson et al., 2012). The numbers indicate how deeply the service would have penetrated to the daily lives of people.

Its astonishing prevalence and the influence over people’s various aspects of lives, including what to believe, how to act, and how to behave, led research on Facebook to span across diverse topics; motivations and behaviors (Acar, 2008; Grimmelmann, 2009; McLaughlin & Vitak, 2012; Moore & McElroy, 2012; Park et al., 2009; Ross et al., 2009; Sheldon et al., 2011), identity presentation (Lamp et al., 2007; Park et al., 2011; Walther et al., 2009), role in

social interactions (Bakshy et al., 2012; Craig & Wright, 2012; Ellison et al., 2007; Ellison et al., 2010; Harlow, 2012; Tong et al., 2008), privacy issues (Acquisti & Gross, 2006; Grimmelmann, 2009), and critical approach (Bucher, 2012; Dubrofsky, 2011; van Dijck, 2012).

The very nature of how Facebook first started its service adds up to its distinctiveness. Unlike other SNSs, Facebook began in 2004 as a Harvard-only service; access was given to people only with Harvard University e-mail accounts. Facebook expanded its users to other university networks, still requiring university e-mail accounts for verification, then to high school students and professionals with corporate e-mails in 2005. Finally in 2006, Facebook signup became available to everyone, yet, still maintaining exclusive networks of universities and corporates. Even though anyone could use the service and develop relationships with random people first met online, this SNS service features a community grounded in real world (Levinson, 2009). Studies support such reality-based relationships on Facebook, providing empirical evidences that people generally use Facebook in order to maintain relationships that are already formed offline (Ellison et al., 2007; Lampe et al., 2006; Sheldon, 2008).

Even though majority of the ‘friends’ on Facebook are based on offline-based relationships, recent studies show that the relationships include broader spectrum of relationships ranging from complete

strangers to intimate relations (Ellison et al., 2010; Hampton et al., 2011; Marwick & boyd, 2010). Therefore, Facebook ‘friends’ should be studied distinctively from the real-world friends.

Using Facebook for investigating people’s communication behaviors may help overcome the previous methodological limitations in regard to observation and self-reporting. Partly due to the extension of offline-based relationships and shared networks, user profiles filed on Facebook tend to be more honest and trustful than other SNSs like Twitter (Lampe et al., 2007; Wilson et al., 2012). When people share social networks or have mutual friends who could verify information on other persons, they show less deceptive self-presentation online and tend to show who they “really” are (Berger, 1987; Donath & boyd, 2004). Thus, the general relationship formations on Facebook may require people to put up a rather trustful profile statement of the user, making Facebook more appropriate for academic research—as in case of this current study—where researchers rely on the self-presented information of the users.

1) Facebook User Information

Facebook user profile contains information on a number of personal attributes. Whether to make the information public or not, determined by the privacy level that a user sets, depends on the user. Yet, there are several basic profile elements that are always required to

be publicly available; name, profile picture, cover photo, sex, and networks (if saved to the system).² The openness of other contents depends on the user. Setting choices are “public,” “friends,” “only me” in general, and besides the three, specific setting could be done in “custom.” In detail, sections of personal attributes on Facebook profile includes sex, age, ethnicity, interests, relationship status, religion, political views, networks, location, school, workplace, and social groups of the user and so on. The importance of sharing such attributes and its influence on boosting interactions among ‘friends’ has been emphasized, especially those of latent ties (Xiang et al., 2010).

2) Interaction Behaviors in Facebook

Moreover, Facebook’s activity logs provide a useful way to study interactions between ‘friends.’ Earlier studies did not take the variety of communication types available on Facebook into consideration and mostly measured Facebook usage through items such as “In the past week, on average, approximately how many minutes per day have you spent on Facebook?” (Ellison et al., 2007) However, measuring Facebook usage through such broad questions may overlook Facebook’s uniqueness of diverse communication channels, which are

² Data Use Policy, Facebook Inc. Available at: <http://www.facebook.com/about/privacy/your-info>

actively utilized by users on Facebook in order to interact with their ‘friends.’ Even Facebook, the company itself realizes the importance of the amount and nature of interaction types. Contents shown in ‘News Feed’ page, which is the primary feature of Facebook, are determined by an algorithm based on the assumption that some types of interactions are considered more important than others, and thus, interaction types and frequency differ according to the intimate level between ‘friends’ (Bucher, 2012). Activities between ‘friends’ on Facebook include “poking” each other, pressing “like” buttons on posts (status updates, specific comments on certain status updates, photos, specific comments on photos, links, comments on links, videos, comments on videos) uploaded by the user, ‘friends’ or where the user or ‘friends’ are “tagged” in, writing “comments” on such posts and messages—classified as “comments”— and “tagging” each other on a photo or in “comments.”

Recently, Ross and his colleagues (2009) conducted a research on the influences of personality on the usage in more detail by incorporating many interaction types. They included posting photos of others, “commenting” on others’ photos, posting on others’ “wall” posts, sending messages, “poking” to the frequency of Facebook use. Also, Moore and McElroy (2012) studied the influences of personality on the time, frequency, and regrets on Facebook usage. Interaction

types of “commenting” on others’ photos, posting on others’ “wall” posts were included. Both research made attempts to specify different communication means that are particular in Facebook through conducting surveys. Yet, self-reporting of usage behaviors still has limitations, especially on recalling the respondents’ actual behaviors and in that the data cannot really show interaction behavior between ‘friends’ since it requires more than one party, yet only the respondent is reporting his/her behavior.

Few studies on Facebook actually collected activity log data in order to study interaction behaviors of Facebook users, overcoming such limitations of survey. In a study of developing a predicting model of tie strength between ‘friends’ by various information available on Facebook, Gilbert and Karahalios (2009) incorporated various aspects of Facebook usage such as “days since last communication,” “distance between hometowns,” “inbox intimacy words,” and specific interaction types such as “participant-initiated wall posts,” “inbox messages exchanged,” “participant’s status updates,” “friend’s photo comments,” “appearances together in photo,” and “links exchanged by wall post.” Moreover, in a more recent research of modeling the relationship strength as a hidden cause of user interactions, Xiang et al. (2010) studied “commenting” on “wall” posts and “tagging” in photos. Even though the research provided

many insights and interesting results, the most common and active interaction types on Facebook such as “liking” photos, “commenting” on photos, “liking” statuses were not investigated.

3. Limitations on the Literature

In summary, despite the necessity to understand communication behaviors of people in the online context as people are “always on,” most of the research on homophily has been yet focused in the offline context. While recent studies have been making attempts in order to investigate the phenomenon online, there are still two main limitations on the research.

First of all, including the recent study of Thelwall (2009), the literature shows that studies on homophily have been conducted in incomprehensive ways. By investigating single attributes separately when identifying homophily effects such as age homophily and geographic homophily, previous studies have not measured the effect of relative attribute similarity effects on homophily considering multiple attributes altogether; the necessity of comparative measure of similarities has been overlooked. Such measure is necessary because it is not a single attribute that influences relationships, but multiple characteristics varying in the similarity.

Kossinets and Watts (2009) made an attempt to investigate homophily phenomenon by its origin, whether it is from personal preferences on attribute similarities or from structural propinquity; thus, measured total similarities that affect ‘choice homophily,’ and separately, total similarities that affect ‘induced homophily.’ Since the goal was to identify the origin of homophily effect, not the difference in the relative amount of each attribute’s influence on homophily, they adopted a simple aggregate similarity measure for ease of interpretation. This aggregates the total similarity by providing equal weight on each attribute and simply adding them up; counting the number of shared attributes. For example, when measuring the total similarity of gender, age, status, and field, the aggregate similarity would range from 0 to 4.

Nevertheless, depending on specific situations in which people are put, particular attributes become prominent among multiple layers of their “social identities”—defined as group membership in categories as age, sex, religion, and ethnicity (Steele et al., 2002). Hence, adding up the number of common attributes in order to measure similarities between people disregards contexts in which people are situated. Accordingly, homophily effects of personal attributes, which were extensively studied within offline contexts should be carefully examined and be applied to the online context. Further, even within

the online context, difference in the effect of attributes may exist among blogs, SNSs, online shopping sites and so on since each service or website possesses distinctive characteristics.

Thus, the first research question is established in order to explore which personal attributes have prominent effects on homophily and how prominent relative to other attributes in the online context, particularly in Facebook; in purpose of providing a more comprehensive and comparative approach towards similarity measurement among diverse personal attributes.

Research Question 1: How do similarities on each personal attribute—listed in Facebook user profile—correlate to the intimacy level of the user?

Moreover, though recent studies on homophily in online contexts empirically study the phenomenon by investigating a positive correlation between similarities and ‘interaction behaviors,’ the initial coinage of the term homophily does not exactly refer to the tendency of similar people communicating more with each other than less similar one. Rather than the observational communication behavior itself, the effect originally indicated “attachment” to similar others, which is a perceptual conception. However, interaction behaviors are indeed important in that without social contact or interaction, relationships cannot be formed or maintained at the first place no

matter how intimate or superficial, and therefore, empirical research on the dynamic process of relationship formation—and maintenance to dissolution—through observing communication between people is critical.

While Lazarsfeld and Merton (1954) did emphasize that relations such as friendships are undoubtedly observed patterns resulting from social interaction between people, the validity to positively correlate similarities directly to interaction behaviors has not been proved empirically. In online contexts, since diverse forms of interaction between people are possible, homophily phenomenon has been observed not only through general communication behaviors between people, but through various kinds of interactions such as information diffusion and e-mail exchange (Choudhury et al., 2010; Kossinets & Watts, 2009).

Understanding homophily by interaction is particularly important for studies that utilize computer data log, other than interviews or survey questionnaires in that large discrepancies were shown between self-reports and actual observations on interactions in many studies (Marsden, 1990). By solely relying on the data, it is virtually unfeasible to know whom one feels more attached or intimate to. Perhaps that partially explains why recent studies conducted by collecting log data observed homophily directly through interaction

behaviors without taking intimacy or emotional closeness in consideration. Since such approach is fundamentally based on the assumption that interaction behaviors reflect perceived intimacy, in advance, it is necessary to investigate, empirically, whether the approach of examining homophily effect by the direct correlation between similarity and interaction level is applicable in online contexts.

Earlier literature was primarily focused on face-to-face communication, where somewhat limited types of interaction behaviors were available than in the online context of Facebook, and thus, it lacks theoretical expectations. To provide exploratory findings on the relation between intimacy and diverse interaction behaviors, the second research question is established as the following:

Research Question 2: How do communication behaviors in various interaction types—available on Facebook—correlate to the intimacy level of the user?

III. RESEARCH METHOD

Data for the research were collected in two ways, through survey and data crawling. In order to measure intimacy levels that users feel towards each of their “friend,” a survey was conducted. Also, to measure similarities and interaction levels between people, users’

profiles and activity log data (between users and each of their “friend”) were crawled through Facebook API.

1. Issues on Facebook User Data Collection

Despite a number of studies on Facebook, only a few works conducted research on homophily; focusing on homophily within a university directed by attributes that inevitably increase propinquity such as year of class, registered courses, and group activities (Clouston et al., 2009; La Fond & Neville, 2010). Absence of studies probing empirical evidence of homophily on Facebook based on more diverse attributes partly results from a methodological issue due to technical constraints.

On account of quite strict emphasis on privacy concerns, data collection through the Open ‘Application Programming Interface (API)’ on Facebook is more difficult and requires extra efforts, than on other SNSs such as Twitter.³ While API could be ‘crawled’ in automated means in Twitter, because user data on Facebook are protected, it is necessary to acquire authorization from each user in

³ Open API (often referred to as OpenAPI new technology) is a word used to describe sets of technologies that enable websites to interact with each other by using REST, SOAP, JavaScript and other web technologies. While its possibilities aren't limited to web-based applications, it's becoming an increasing trend in so-called Web 2.0 applications. The term API stands for Application Programming Interface. (Open API, Wikipedia. Available at: http://en.wikipedia.org/wiki/Open_API)

order to collect the data.⁴ As ways to cope with the restriction, research on Facebook generally recruits participants through voluntarily in offline contexts, via Facebook Applications (Apps), or the combination of the two (Acquisti & Gross, 2006; Nazir et al., 2008).

A specifically designed way to utilize Application Programming Interface (API) in order to crawl user data is to develop a Facebook Application (App) that utilizes the API: Procedures on creating Facebook App are well explained on the App “Developers” page: Procedures on creating Facebook App are well explained on the App “Developers” page.⁵ Then, by requesting and acquiring permission on data access to the App users, user information and activity logs will be accessible. Authorization of the App is required in order to gain access to all user information available. Generally, OAuth Dialog for authorization is recommended in that it allows application developers’ access to users’ personal information without knowing passwords of each user. By redirecting user’s browser to the “request for permission” dialog, the OAuth authorization is processed. Basically, the user is only asked to authorize the App to access basic information

⁴ A Web crawler is a computer program that browses the World Wide Web in a methodical, automated manner or in an orderly fashion. Other terms for Web crawlers are ants, automatic indexers, bots, Web spiders, Web robots, or—especially in the FOAF community—Web scutters. This process is called Web crawling or spidering. (Web crawler, Wikipedia. Available at: http://en.wikipedia.org/wiki/Web_crawler)

⁵ Facebook Developers, Facebook Inc. Available at: <https://developers.facebook.com/apps>

that is available publicly or by default on Facebook. Thus, in order to collect additional information on users, developers must request for further permissions for authorization to each of the user.⁶

Looking into the homophily phenomenon in a comprehensive way without actual data on users' profiles and their activity logs seems practically meaningless, since homophily research requires to analyze information on users' personal attributes and association in their networks altogether. Despite the unavailability of an "easy" crawling process, in order to account for homophily online, a Facebook App was developed for current research. Not to mention user profiles and their behavioral information, real-time updates on users' activities on Facebook will be also available to the App developers.⁷

Considering the fact that initial Facebook user pool was recruited throughout affiliated networks and that the service primarily revolved around such networks, including all kinds of affiliation of the user seems necessary, if available. Some argue that Facebook contains a demographic bias in nature in that Facebook started its service based on users within university networks (Hargittai, 2007; Lampe et al., 2007; recited from Papacharissi, 2009). However, data on the demographic bias in Facebook cannot be measured in current study,

⁶ For specific descriptions on the entire accessible data, refer to Appendix 1.

⁷ Graph API, API Reference, Facebook Inc. Available at:
<http://developers.facebook.com/docs/reference/api/>

because only profile information and activity logs of the Facebook App users are accessible through the App. This implies the difficulty to calculate the potential demographic distribution from the entire Facebook user pool; that is, it is difficult to rule out demographic bias from the homophily effect in Facebook.

2. Participant Recruitment

Participants for the experiment were recruited through several online channels. Since current research is to explore homophily effect specifically in Facebook, participants were limited to Facebook users. Thus, a brief explanation on the study, with a link leading to the experiment website and a quote asking to share the post, was initially uploaded and promoted on Facebook by the Facebook account of current researcher throughout the ‘friends.’ Additionally, messages to recruit Facebook users were posted in popular online websites and campus community sites.

3. Survey

A web application using Facebook API was specifically built for the survey. The application was built with Ruby on Rails and was designed to collect participants’ responses and their actual Facebook

activity data.⁸ When participants sign in to the application with their Facebook account, agreement on collection of their profile information and activity log data were requested by the App; the survey was initiated only for those who agreed to the request. Once participants agreed to the data collection, they were asked to rate intimacy levels they feel towards each of the 27 randomly selected ‘friends’ of theirs and 3 fake users so that unreliable participants could be excluded from the study; 3 fake users were randomly selected among a list created of 10 fake users [Figure 1].⁹

hci-d lab. Signed in as Jieun C. Wee | Sign out

페이스북 '친구'에게 느끼는 친밀도 조사

Facebook Study (1/7)

다음의 질문과 관련하여 제시된 페이스북 친구를 평가해 주세요.

1. 당신은 다음의 친구와 얼마나 가까운 사이입니까?
(0: 모르는 사람, 1: 전혀 가깝지 않음 -- 7: 매우 가깝음)

	(0) <input type="radio"/> (1) <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> (7)
곽기욱	
	(0) <input type="radio"/> (1) <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> (7)
Teo Kim	
	(0) <input type="radio"/> (1) <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> (7)
Seyong Ha	
	(0) <input type="radio"/> (1) <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> (7)
Seokhyun Nate Yee	
	(0) <input type="radio"/> (1) <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> <input type="radio"/> (7)
Young Ho Byun	

Figure 1. Example on the survey questionnaire

⁸ Ruby on Rails is an open-source web development framework. (<http://www.rubyonrails.org>)

⁹ [Figure 1] is a partial list of 27 randomly suggested ‘friends’ and 3 fake ‘friends,’ generated for the researcher. The same subset of 30 ‘friends’ list is rated across 7 items to measure intimacy towards each of the ‘friend.’

Same subset of ‘friends’ and strangers was suggested across 7 different items, which were all measured in 7-point Likert scale with 1 for the least intimate, 7 for the most intimate, and 0 for “don’t know the person.” 7 items to rate the overall level of intimacy on each ‘friend’ was adopted from the study of Vangelisti and Caughlin (1997), in which the items had alpha reliabilities of .93 in explaining psychological closeness. Specific questionnaires include “How close are you to your ‘friend’?”, “How much do you like your ‘friend’?”, “How often do you talk about personal things with your ‘friend’?”, “How important is your ‘friend’s’ opinion to you?”, “How satisfied are you with your relationship with your ‘friend’?”, “How much do you enjoy spending time with your ‘friend’?”, and “How important is your relationship with your ‘friend’?”.

4. Experiment Procedure

As participants sign in with their Facebook account, a request on access to personal information and activity log data on Facebook appeared [Figure 2].¹⁰ Once the permission was granted by the participant, the experiment website (web application created for the survey) accessed data of the participant’s Facebook user account via OAuth authentication provided by Facebook. OAuth is an

¹⁰ Refer to Appendix 1 for the full list of permission requested to Facebook in order to access listed user data

authentication method that is commonly used in online applications, which enables applications to not only verify users but also access information on user profile and activity log data on Facebook through the account (Suh et al., 2012).

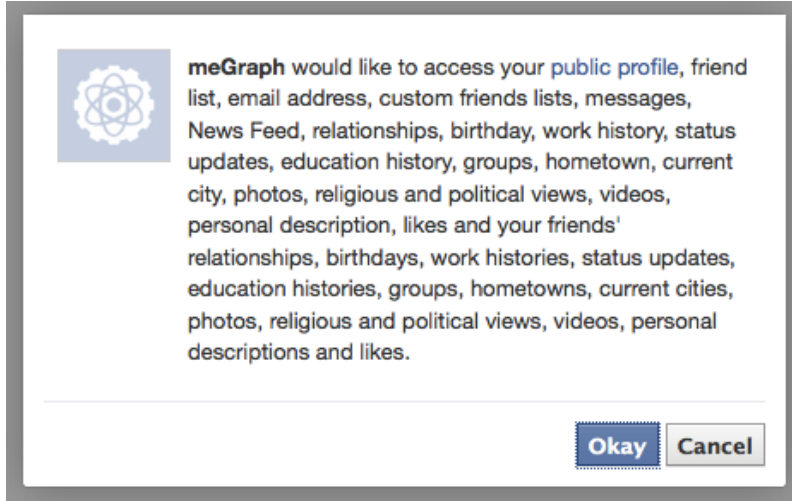


Figure 2. Request on access to participant's Facebook data

Then, the participant was instructed to begin a survey on 30 of their randomly selected Facebook 'friends,' while as noted earlier, the 3 among the 30 were fake 'friends' included in order to exclude participants with unreliable responses. The subset of 27 'friends' was extracted randomly among participants' Facebook 'friends' who had interactions with the participant in 50 recent posts. 3 fake 'friends' were randomly selected from the 10 fake 'friend' list. All of the evaluated 'friends' were listed with their Facebook account names and profile pictures so that participants could better identify their 'friends'

[Figure 1]. Each of the ‘friends’ was rated in a 7-point Likert scale according to the questionnaire.

The experiment website (Facebook web application) was open to public for 5 days from June 3rd, 2013 to June 8th, 2013 and total number of 75 Facebook users participated in the study. Among the 75 participants, those who failed to identify either one of 3 fake ‘friends’ as 0, controlled their privacy settings so that all Facebook applications were blocked from their data, or did not finish the entire survey were excluded from the analysis. In cases where participants actually had real ‘friends’ with the same name as the fake ‘friend,’ and therefore, rated certain ‘friend’ as other than 0 were not excluded. The final number of participants included in the study was 36, resulting in a dataset of 972 ($= 36 \times 27$) rated Facebook friendships for analysis. However, one of the 36th participant’s ‘friends’ prevented access from Facebook applications, and due to that, Facebook data of that specific user profile and interaction log between the user and the 36th participant was not collectable. Therefore, the final dataset contained 971 Facebook friendships.

5. Data Crawling

1) Similarity

Similarity indicates sharing personal attributes between a participant and a ‘friend.’ In order to measure similarities in Research Question 1, data on Facebook user profiles of all 36 participants and subsets of their ‘friends’ (27 for 35 participants, and 26 for 1 participant) were collected as Table 1. While current study initially planned to include both political views and religion in addition to attributes classified in Table 1, they were excluded in the analysis because most of the users did not list information on the two sections.

In case of work and education, additional speculation was required because there were a number of different ways to register for the identical information. For example, in terms of college, Seoul National University was described diversely as “서울대학교,” “SNU,” “Seoul National University (서울대학교),” and such information were treated as different even though they all referred to the same institute. Thus, the researcher went over the collected data on personal attributes of work and education in order to check for mismeasurement regarding such labeling issue.

Table 1. Similarity variables

Attribute	Variable	Description
sex	co_gender	accordance of gender between a subject and a friend
age	age_difference_by_year	age difference between a subject and a friend by the year born
location	co_location	accordance of current residence (city/country) between a subject and a friend
hometown	co_hometown	accordance of the hometown (city/country) between a subject and a friend
work	co_work(r)	rate of the number of same workplaces of a subject and a friend relative to the subject's registered workplaces
education	co_highschool(r)	rate of the number of same high schools of a subject and a friend relative to the subject's registered high schools
	co_college(r)	rate of the number of same colleges of a subject and a friend relative to the subject's registered colleges
	co_gradschool(r)	rate of the number of same graduate schools of a subject and a friend relative to the subject's registered graduate schools
	co_concentration(r)	rate of the number of same concentration of a subject and a friend relative to the subject's registered concentrations
groups	co_groups	number of same groups of a subject and a friend
interests	co_liked_pages	number of same liked pages of a subject and a friend
marriage status	co_relationship_status	accordance of relationship status between a subject and a friend
network	mutual_friends(r)	rate of the number of mutual friends of a subject and a friend relative to the number of the subject's friends

2) Interaction

Interaction refers to activities that occur between users on Facebook. Since articulated networks on Facebook as ‘friends’ do not necessarily mean all of the ‘friends’ have the same tie strength, and the communication frequency itself does not directly reflect the affiliation between people, another criterion indicating some level of intimacy needs be included in order to measure homophily effect through interaction behaviors (boyd & Crawford, 2012). The idea is based on the assumption that people will choose different interaction types to communicate with others depending on the level of intimacy towards the partner. Even in the study of Wellman and Wortley (1990), the frequency of telephone contact showed significant correlation with the strength of relationship, whereas the frequency of face-to-face contact did not.

People tend to allocate resources such as time to form or maintain relationships that they regard more important since the resources are limited (Dindia & Canary, 1993; recited from Xiang et al., 2010). Also, in a recent work of examining strength of relationships in SNSs, researchers underlined that the type of interaction and the likelihood vary according to relationships strengths, and thus, different weight should be given on each interaction type in terms of intimacy (Xiang et al., 2010; recited from Wilson et al., 2012).

Thus, every post and activity type will be considered in the study as Table 2 and Table 3 for all 971 relationships, which are the relationships between 27 ‘friends’ for each of 35 participants and 26 ‘friends’ for 1 participant. Interaction behaviors that occurred on timeline were collected based on 300 recent posts uploaded since April 1st, 2012 [Table 2].¹¹ Total amount of data collected and analyzed on each activity type were 8,491 “comments,” 16,171 “likes,” 2,542 “tags,” and 47,348 “chat” for all 971 relationships.

Unlike interactions on timeline, message data (inbox/outbox) containing both synchronous ‘chatting’ and asynchronous ‘messaging’ are only collectable in message box threads based on the subjects of each chatting [Table 3]. Also, unlike posts uploaded on timeline, messaging only includes one specific activity type; send and receive “comments.” Thus, 50 recent message boxes and individual messages within the thread of each of all 36 participants were crawled. Then, among the data, message exchanges only with the suggested ‘friends’ of each participant were separately collected for analysis.

¹¹ An activity type of “poke” was not included in the study because all “poke” messages are stored in encrypted form and retained for two days after the last recipient views the poke — a process that helps facilitate abuse reporting. Since the encryption key of the “poke” is deleted after that period, collecting data on “poke” for interaction during the whole period was unfeasible, thus, excluded.

(Available at: <http://techcrunch.com/2012/12/22/your-facebook-pokes-are-stored-for-two-days-then-their-encryption-keys-are-deleted/>)

Table 2. Interactions on timeline

Post Type	Description
link	posts with links
video	videos uploaded by the user and where the user is tagged ('videos/uploaded' + 'videos')
photo	photos uploaded by the user and where the user is tagged ('photos/uploaded' + 'photos')
status	status updates without links, videos, and photos

Activity Type	Description
likes	'likes' directly on all the post type (post_likes) + 'likes' on a comment of all the post type (comment_likes)
comments	'comments' directly on the post type (post_comments) + 'tagged' on a comment of the post type (comment_tags)
tags	'tagged' directly on the post

Table 3. Interaction on messenger (Chat / Message)

Post Type	Activity Type	Description
chat	chat_comments	exchanged chat messages

Interaction types are specified into the combination of post types and activity types, yet primarily based on the activity type. For analysis, while the activity of "tags" is specified into post types, "likes"

and “comments” activity types incorporate all post types. The rationale for the classification is that “tags” vary in the form according to post types. For example, “tags” on photos are often done directly to the photo, identifying individuals by the tags, sharing a photo taken together or showing that they are/were together. On the other hand, “tags” on statuses are often used as sending direct message to certain people in a more public way or telling people that they are/were together. Moreover, “tags” on photos are presented on the photo while on statuses are only presented as texts.

IV. RESULTS

1. Descriptive Analysis

Descriptive analysis on the participants shows that there were twice more females than males (24:12) [Table 4]. Age distribution was highly skewed towards the 20s, indicating that the result of current study reflects homophily in Facebook of the young. 75% of the participants referred themselves as “always on” on Facebook, presenting how deep Facebook has penetrated into their daily life. 42% of participants responded that they spent 10 to 30 minutes a day on Facebook. The result of Facebook usage enables complementary interpretation: Because participants are connected to Facebook 24/7,

they have no need to stay in Facebook page for a long time, and rather, they respond to notifications on Facebook instantly and right away.

Table 4. Basic information of participants

Characteristics		Male	Female
Age Distribution	19 – 25	9	17
	26 – 30	3	6
	above 30	0	1
Frequency of Facebook Use	“don’t use every day”	0	0
	1 – 2 times a day	1	0
	more than 3 times a day	2	6
	“always on”	9	18
Daily Time Spenton Facebook	less than 10 min	0	1
	10 – 30 min	5	10
	30 min. – 60 min	2	4
	60 min. – 120 min	1	8
	120 min. – 180 min	2	1
	more than 180 min	2	0

2. Multiple Linear Regression (MLR)

Correlations between the survey data (intimacy) rated by participants and the data collected via the experiment website (similarity and interaction) are analyzed through Multiple Linear Regression Model. Intimacy, evaluated by 35 participants towards

each of their 27 ‘friends’ and a participant towards 26 ‘friends,’ is taken as the dependent variable for the study. In all 971 relationships, the intimacy score was measured as the mean score of 7 survey items. How similarities as independent variables, and further, how interaction types as independent variables correlate to intimacy are explored, respectively, in following sections.

1) Similarity on Attributes that Composes Intimacy

Similarity is measured by shared attributes between a participant and a ‘friend.’ Attributes that compose similarity are taken into consideration based on Facebook user profile information of each user. Since how much information the user put up on the profile page fully depends on the user’s decision, there were missing values across attributes. In a study of Lampe and colleagues (2007), it has been found that on average, users complete 59% of the fields available to them. Especially, only a small number of users listed their political views or religion, resulting in missing values of the shared values on them; thus, the two attributes were excluded in the study.

Missing values (NA) in other attributes were treated as 0. Even though Facebook networks are primarily based on offline relationships, the conceptual range of “friendship” in Facebook is broader, including complete strangers. Hence, when information on certain attribute of the user such as workplace is not exhibited on the profile page of a

‘friend,’ then there is no similarity on workplace that the user recognizes toward the ‘friend.’ Other than missing values, there was an additional case to handle for attributes—work, education, and networks—that were calculated at the ratio. When the total number of a participant’s registered profile on certain attribute is 0, ratios are fundamentally impossible because the denominator would be 0 then [Table 1]. Cases relevant to such issue were specified as “NaN,” and were replaced with 0 as in the case of “NA,” for the same reason.

Table 5. MLR analysis result: Similarity

Similar User Attributes (Independent Variable)	β	σ	p	<i>Sig.</i>
(Intercept)	3.607	0.138	<0.001	***
co_gender	0.146	0.13	0.26	
age_difference_by_year	-0.022	0.019	0.255	
co_location	0.116	0.149	0.436	
co_hometown	-0.214	0.175	0.221	
co_work(r)	0.624	0.472	0.187	
co_highschool(r)	-0.044	0.184	0.81	
co_college(r)	-0.242	0.171	0.157	
co_gradschool(r)	0.632	0.361	0.08	.
co_concentration(r)	-0.157	0.264	0.552	
co_groups	0.842	0.635	0.185	
co_liked_pages	0.004	0.006	0.47	
co_relationship_status	-0.326	0.205	0.112	
mutual_friends(r)	0.55	0.608	0.365	

.p < .1 *p < .05, **p < .01, ***p < .001

Results of multiple linear regression analyses on similarity (independent variable) and intimacy (dependent variable) is summarized in Table 5. Unlike the literature on homophily that has been mostly studied in offline contexts, no significant correlations were found between attribute similarities and intimacy. While homophily effects on certain attributes were found in the study of Thelwall (2009), only the similarity on graduate school showed statistically, yet, a low significant correlation to intimacy.

2) Interaction Types that Predict Intimacy

In advance of exploring how each interaction type predicts intimacy, a plot was drawn in order to see how the simply counted interactions (without considering type differences) were correlated with intimacy scores. Since the overall interaction counts had great variance, interaction variable was logarithmically transformed after adding 1 to interaction. In Figure 3, each dot represents a relationship; total 971 dots in the plot. Even though the plot shows linear correlation between interaction frequency and intimacy, residuals on each case vary considerably.

The overall aggregated interaction behaviors of 36 participants are descriptively specified in Table 6. While current study classifies interaction types in a more activity-based combined way as in Table 7,

descriptive analysis on the interaction behavior was conducted in more detail.

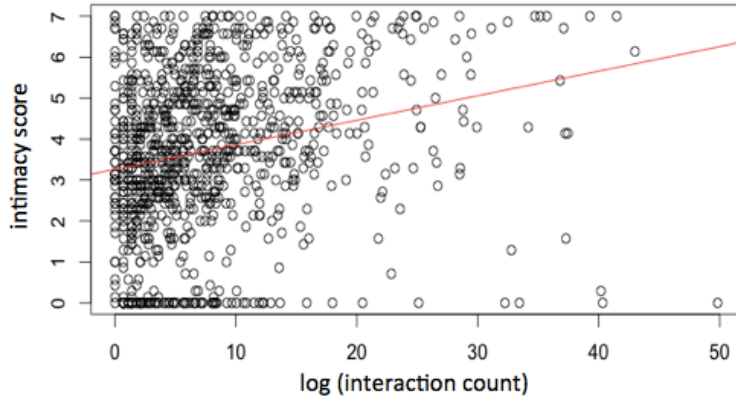


Figure 3. Plot on intimacy and logarithmic transformation on interaction (simple count) of 971 relationships

Overall, link and video posts are seldom used as a communication channel between participants and ‘friends.’ In particular, “tags” activities on link and video types both never occurred among the participants’ relationships in current study. Though message post type occurred most frequently, the standard deviation of the 36 participants’ messaging activities were 338.826, which is far greater than the mean. This indicates that there were participants who actively used messages with their ‘friends,’ and those who nearly used messages to interact with their ‘friends.’ “Likes” appeared to be the most popular way to communicate regardless of post types, followed by “comments.” Among many different possible communication channels on Facebook,

“likes” and “comments” on photos and statuses were the most actively used interaction types on Facebook.

Table 6. Descriptive analysis on interactions of 36 participants

Post Type	Total Frequencies	Mean	SD
link	2,063	0.424	2.309
video	29	0.006	0.089
photo	13,235	2.723	9.824
status	13,496	2.777	8.201
chat	54,829	56.408	338.826
Activity Type	Total Frequencies	Mean	SD
post_likes	15,952	4.102	11.108
comment_likes	8,880	2.284	8.709
post_comments	671	0.176	1.385
comment_tags	1,229	0.316	1.386
tags	2,091	0.538	2.231
Post_Activity Type	Total Frequencies	Mean	SD
link_post_likes	1,292	1.329	4.536
link_post_comments	560	0.576	2.04
link_tags	0	0	0
link_comment_likes	98	0.101	0.587
link_comment_tags	113	0.116	0.616
video_post_likes	0	0	0
video_post_comments	20	0.021	0.169
video_tags	0	0	0
video_comment_likes	4	0.004	0.078
video_comment_tags	5	0.005	0.072

photo_post_likes	7,940	8.169	16.613
photo_post_comments	3,453	3.552	12.119
photo_tags	588	0.0605	2.698
photo_comment_likes	465	0.478	1.541
photo_comment_tags	789	0.812	2.617
status_post_likes	6,720	6.914	12.181
status_post_comments	4,847	4.987	11.646
status_tags	83	0.085	0.377
status_comment_likes	662	0.681	2.16
status_comment_tags	1,184	1.218	3.418

After descriptively looking into the general picture of users' interaction behaviors on Facebook, multiple linear regression analysis was conducted to identify interaction types that predict intimacy between participants and their 'friends.' As explained in Figure 3 and Table 6, because the interaction frequencies varied considerably across 36 participants and across 971 relationships, logarithmic transformation on interaction was applied after adding 1 to all of the frequencies of interaction types [Table 7].

Interaction types were measured by dividing each activity type—each additionally separated by cases where a participant was the sender (*to_friends*) and where a participant was the receiver

(from_friends)—into total frequencies of the activities (regardless of “from” and “to”).

Table 7. MLR analysis result: Interaction (linear-log model)

Interaction Types (Independent Variable)	β	σ	p	Sig.
(Intercept)	3.556	0.07477	<0.001	***
likes_from_friends	-8.919	4.486	0.047	*
likes_to_friends	7.922	4.319	0.067	.
comments_from_friends	15.913	5.833	0.006	**
comments_to_friends	6.372	5.626	0.258	
photo_tags_from_friends	-15.661	17.214	0.363	
photo_tags_to_friends	-28.91	16.824	0.086	.
status_tags_from_friends	102.783	67.5	0.128	
status_tags_to_friends	204.187	166.45	0.22	
video_tags_from_friends	NA	NA	NA	NA
video_tags_to_friends	NA	NA	NA	NA
link_tags_from_friends	NA	NA	NA	NA
link_tags_to_friends	NA	NA	NA	NA
chat_from_friends	4.289	1.851	0.021	*
chat_to_friends	-3.885	2.019	0.055	.

.p < .1 *p < .05, **p < .01, ***p < .001

For example:

$$\text{likes_from_friends} = \frac{\begin{array}{l} \text{\# of “likes” a participant} \\ \text{received from a ‘friend’} \end{array}}{\begin{array}{l} (\text{\# of “likes” a participant sent to a ‘friend’} + \\ \text{\# of “likes” a participant received from a ‘friend’}) \end{array}}$$

“Comments”

The most significant interaction type turned out to be `comments_from_friends`, with a high beta coefficient, while `comments_to_friends` did not show any statistic significance on intimacy and had less than half of the beta coefficient. The result indicated that participants felt more intimate to ‘friends’ who made “comments” frequently to the participant. Moreover, “comments” made from the participant to a ‘friend’ did not predict intimacy the participant felt towards the ‘friend.’ This could be put in another way: Though a user tend to feel more intimate to ‘friends’ who make “comments” to the user, the user does not necessarily make more “comments” to ‘friends’ that the user feel more intimate to.

“Likes”

“Likes” activity provides an interesting correlation with intimacy. While `likes_from_friends` had significance with a negative beta coefficient, `likes_to_friends` nearly had significance with a positive beta coefficient. The result may seem contradictory to “comments” results.

“Chat”

“Chat” activity on the messenger shows an interesting finding as well. `Chat_from_friends` has a significant correlation to intimacy, with a positive beta coefficient. The significance indicates that the

more a ‘friend’ sends “chat” to a participant, the more intimate the participant would feel towards the ‘friend’; similar to the case of `comments_from_friends`. `Chat_to_friends` showed negative correlation to intimacy even though the significance was low. Interpretations on the interaction type of sending “chat” messages from and to ‘friends’ required further elaboration.

Others

The NA results on “tags” of video and link posts are due to the 0 value of the frequency [Table 6]. The zero frequency of interaction through videos and links itself may provide a critical point of communication behaviors related to intimacy; both of the post types are variables irrelevant in modeling the prediction of intimacy. Yet, since there were activities as “likes” and “comments” on the both links and videos, jumping into conclusion of excluding them was not reasonable. Thus, both of the variables were included in the analysis.

V. DISCUSSION

Main findings from current research are related to measurement issues of homophily effect in the computer-mediated context of Facebook. First of all, the traditional way of observing homophily phenomenon as affiliation of those who share similar attributes turned out to be not applicable to the online study. In case of Facebook,

similarity on user profile showed no significant level of correlations with intimacy.

One possible interpretation on the result is that in offline contexts, people actually “see” and “meet” each other, and thus, are able to confirm or feel the similarity between each other. In the study of Ellison et al. (2010), it has been discovered that the social information seeking behaviors of people on Facebook only affect connection strategies with their ‘friends’ in case of latent ties, but do not influence close friends. On the other hand, in online contexts as Facebook, people often get cues or get to know about other people through their profile pages. Considering the wide spectrum of “friendship” Facebook covers, presumably, people would rely heavily on profile pages of their ‘friends’ in order to get to know or keep up with them, since people have limited time to actually “catch up” personally with all of them in interest. Yet, contents on profile pages are registered by ‘friends’ themselves, and therefore, there is high possibility that people would not give much credit or value to the information that are “exhibited” on Facebook.

Taking the perspective of Lazarsfeld and Merton (1954), looking into relationship as a dynamic process, keeping relations could be studied in terms of “seeking continued contact,” i.e. interactions. Yet, in the online context such as Facebook, keeping up with someone

could occur in various ways—interaction types. Thus, an alternative approach of studying “online homophily” is suggested. An investigation on correlations of intimacy and diverse interaction types on Facebook was conducted, and certain interaction types appeared to have statistically significant relations to intimacy.

One possible explanation on the somewhat contradictory data between “comments” and “likes” is based on the fact that “likes” activity requires the least effort to engage, but is the most frequently and commonly used channel in Facebook. While generally “likes” do not predict intimacy, too many “likes” from a ‘friend’ could make the participant feel less intimate to the ‘friend’ because the ‘friend’ could be considered as a person who simply “likes” all the posts on Facebook, not particularly the participant’s; otherwise, the ‘friend’ would have left more “comments” to the participant other than pressing the “likes” button every time. Yet, more elaboration on the activity is necessary to make reasonable sense out of the result. Results on photo’s “tags” were somewhat unexpected. Intuitively, if people are tagged together in a photo, or either one of a person “tags” another one, the relationship seems to be close. More speculation on the issue is further necessary, yet, a presumption could be made because there are Facebook applications that publish photos with users’ Facebook ‘friends’ tagged in it; for example, a calendar photo with all

birthdays of Facebook ‘friends’ being tagged in the calendar. To obtain deeper understanding on the result of “chat,” post-interviews on two of the recruited participants were followed up by the analysis. The interviewees were selected based on their intimacy evaluation of ‘friends’ in the survey. No theoretical evidence was available to define and label the range of relationship itself by the intimacy scores. Therefore, among the 36 participants, those who evaluated two or more ‘friends’ with each of the rounded intimacy score of all 1 (the lowest intimacy), 4 (the mid-level intimacy), and 7 (the highest intimacy) were extracted. Such procedure sorted out two participants: interviewee A and interviewee B.

The first interviewee A’s activity log data showed that A exchanged (both received and sent) “chat” messages more often with the more intimate ‘friends.’ However the second interviewee B’s interaction data demonstrated no consistent pattern in the usage of “chat,” especially in case of “chat” messaging with the most intimate ‘friends’; with some of them, B interacted frequently, while with some others, B had no interaction at all. The main purpose of A using Facebook was to keep up-to-date with all the people A wanted to maintain relationships with, however, A felt no need to actually “talk to” or respond to their ‘friends’ who were not so close. The following is a part of A’s interview in the overall usage of Facebook:

“It’s not that I use Facebook particularly with a purpose of keeping in touch with my close friends. I normally check Facebook and see overall News Feeds of my Facebook ‘friends,’ and I just respond to those who I want to talk to or am interested in that moment. Mostly in my case, I don’t really care enough about ‘friends’ who are not really close to me but are just in a same school and is a friend of my friend. I just visit their profile page when I am curious about them. But those who are my really good friends, because we share a lot of our daily lives together, and therefore have many posts in common and stuff, I just talk to them more and respond to their posts more I’m not a fan of Facebook messenger, but it is convenient to briefly talk to my friend when you see them ‘online.’ Mostly, I would call or text them, but when I’m currently using my laptop, it’s just more convenient to talk to my friends through Facebook messenger. But for the non-close friends, I wouldn’t really ‘send’ them a message just because they are “online.”

During the interview, B turned out to have the same motivations of using Facebook as A. Also, despite the different patterns discovered in interaction behavioral data, according to the interview, B actually used Facebook in the same way that A described. The only difference was not based on the interviewee B’s behavior, but due to the difference in Facebook usage of B’s ‘friends’ and A’s. There were a number of ‘friends’ among B’s most intimate ‘friends’ who did not use Facebook “chat” at all, resulting in no “chat” interactions with B on Facebook even though they were very close. Thus, there seems to be a high possibility that among “close” or “confiding” relations, Facebook messenger may be just one of the various personal communication

channels; as the least effective among those because it was not described as the primary channel for one-to-one conversations.

Thus, in case of online contexts, since certain interaction types reflect intimacy, homophily phenomenon could be empirically measured through the interaction behaviors online. However, considering the fact that most of the interaction types appeared to have no significant correlation with the intimacy level, more basic issue could be raised. Especially in ethnographic studies, it has been studied that Facebook is a place where users feel disconnected to their ‘friends’ and most of the friendship on Facebook were established with acquaintances, only because of a social norm that refusing friend request is considered impolite (recited from Tong et al., 2008). Also, it has been found that many Facebook users feel annoyed and hoped their ‘friends’ to keep away from too much personal emotional exposure on Facebook, indicating less emotional closeness to ‘friends’ (McLaughlin & Vitak, 2011). Previous findings from ethnographic studies provide insights in that Facebook could be a space where intimacy is not the key, and rather, expressing and presenting oneself to others in the process of developing and maintaining relationships are in the center of interest.

Also, findings from current study suggest further elaboration on the homophily phenomenon by comparing group of people based on

intimacy; high, middle, and low. Especially, results from the first research question on the non-significant correlation between attribute similarities and intimacy, as well as from the second research question on the interaction type of “chat” enable a presumption on the relationship-based difference in interaction behaviors. As in Marsden’s study (1988), individual attribute homophily effects have high a possibility to differ among wide range of relationships such as superficial relations, close friendships, confiding relations, and kinship. How communication behaviors differ among Facebook users or what role Facebook actually plays as a communication service in people’s relationships according to distinct level of relations could be more precisely inquired through categorizing relationships based on a preliminary study on intimacy scores towards ‘friends’.

For example, findings from “chat” in current research could be understood as Facebook being primarily used as a communication service between acquaintances, with a specific purpose of “keeping up relations” or “checking” with ‘friends’ who are neither close nor a stranger. Those in confiding relations may not care or need to know about ‘friends’ profile on Facebook, because they already know each other well and are intimate enough not to care; moreover, there are many other private or instant communication services available for those in close relationships such as instant messaging, SMS, phone

calls, meeting offline etc. Yet, for those who just got to know each other and are in a process of deciding whether to further develop the relationship into a more intimate level or dissolve it, profile information acquired through Facebook would be critical to their continued affiliation. Such difference in similarity-attraction effect depending on the qualification of relationships has been emphasized in established literature as well (Huston & Levinger, 1978).

Further, there are some issues to be re-considered or improved in current research. First issue concerns the random selection of participants' 'friends' for the survey. Facebook 'friends' of each participant were randomly selected based on two reasons. First of all, selecting 'friends' based on the similarity of personal attributes listed in Facebook user profile was impracticable, since no data on participants or their 'friends' were available beforehand. Also, deciding 'friends' according to the interaction behavior between the participant and the 'friend' on Facebook was fundamentally absurd, because no theoretical expectations or sound evidence were established on the relationship between intimacy and various interaction types used as communication channels on Facebook.

Secondly, the sample size itself is not representative of the Facebook user population. The sample size of 36 had highly skewed age distribution, focused on the 20s. Even though the result cannot be

applied across ages without further speculation, as an exploratory research among Facebook users in their 20s, despite the small sample size, more than 70,000 interactions among 971 Facebook friendships were analyzed. Thus the study provides insights and opportunities to measure homophily from a different perspective that is more relevant to the computer-mediated context. In addition, unlike earlier literature on Facebook usages and interaction behaviors that only investigated the frequency of interaction on the usage, current research looked into interaction behaviors more in detail by distinguishing the sender (from) and receiver (to) of every interaction behaviors between the 971 friendships, and included such factor in analysis (Gilbert & Karahalios, 2009; Moore & McElroy, 2012; Ross et al., 2009; Xiang et al., 2010). Nevertheless, a larger and a more representative sample of the population might have brought more interesting and insightful findings to the work.

Also, causal relationship on the intimacy and interaction could be questionable. This is an inevitable limitation on cross-sectional studies. Especially, considering the social influence process, which occurs when people who associate often tend to become similar, the interaction behaviors may be the cause of the intimacy level, not the result. Such suspicion is a more fundamental question on the homophily phenomenon itself, whether it is the homophily effect or

the social influence process that explains people's relationship better (Xiang et al., 2010). Since this study is focused on the homophily effect itself shown in a computer-mediated context and on better ways to understand homophily in the context—rather than distinguishing the effect from the process of social influence—the question on the causal relationship should be further investigated in future studies based on longitudinal data.

Another issue related to the sample is numerous missing values of user profile information on Facebook. Since such issue is not a factor controllable by researchers, recruiting sample of Facebook users who registered all sections of profile attributes would provide a better understanding on the relationship between user similarities and intimacy. However, the fact that a number of users have unpredictable number of missing information in their profiles makes studying personal attributes relying on the Facebook data difficult. Since such characteristic itself manifests the nature of Facebook, and further, could be the case of other diverse online websites such as SNSs, it is a finding of the current research that alternative indicators other than similarities based on personal attributes should be applied to the homophily phenomenon in online contexts. Based on the study, further investigation on broader range of age cohort and participants with more comprehensive Facebook profile data would incorporate

more representative data of Facebook user population, enabling a generalized explanation on “online homophily.”

VI. APPENDIX

<Appendix 1> Permissions for Data Access

Appendix 1-1. User and Friend Permission

User permission	Friend permission	Description
user_birthday	friends_birthday	Provides access to the birthday with year as the birthday property
user_education_history	friends_education_history	Provides access to education history as the education property
user_groups	friends_groups	Provides access to the list of groups the user is a member of as the groups connection
user_hometown	friends_hometown	Provides access to the user's hometown in the hometown property
user_likes	friends_likes	Provides access to the list of all of the pages the user has liked as the likes connection
user_location	friends_location	Provides access to the user's current location as the location property
user_photos	friends_photos	Provides access to the photos the user has uploaded, and photos the user has been tagged in

User permission	Friend permission	Description
user_relationships	friends_relationships	Provides access to the user's family and personal relationships and relationship status
user_religion_politics	friends_religion_politics	Provides access to the user's religious and political affiliations
user_status	friends_status	Provides access to the user's status messages and checkins. Please see the documentation for the location_post table for information on how this permission may affect retrieval of information about the locations associated with posts.
user_videos	friends_videos	Provides access to the videos the user has uploaded, and videos the user has been tagged in
user_work_history	friends_work_history	Provides access to work history as the work property

Appendix 1-2. Extended Permission

Extended permission	Description
read_mailbox	Provides the ability to read from a user's Facebook Inbox/Outbox

Extended permission	Description
<code>read_stream</code>	Provides access to all the posts in the user's News Feed and enables your application to perform searches against the user's News Feed
<code>read_friendlists</code>	Provides access to any friend lists the user created. All users' friends are provided as part of basic data, this extended permission grants access to the lists of friends a user has created, and should only be requested if your application utilizes lists of friends.

VII. REFERENCE

- Acar, A. (2008). Antecedants and consequences of online social networking behavior: The case of Facebook. *Journal of Website Promotion*, 3(1-2), 62-83.
- Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information, sharing, and privacy on the Facebook. *Privacy Enhancing Technologies*, 4258, 36-58.
- Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. *In Proceedings of the 21st International Conference on World Wide Web*, 519-528. New York: ACM Press.
- Baron, N. S. (2008). *Always on: Language in an online and mobile world*. Oxford: Oxford University Press.
- Berger, C. R. (1987). Planning and scheming: Strategies for initiating relationships. In Burnett, R., McGhee, P., & Clark, D. (eds.) *Accounting for relationships: Explanation, representation, and knowledge*. New York: Methuen, 158-174.
- boyd, d. m. & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13, 210-230.
- boyd, d. & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.
- Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, 14(4), 1-17.
- Burts, R. S. (2000). Decay functions. *Social Networks*, 22(1). 1-28.
- Byrne, D. E. (1971). *The attraction paradigm*. New York: Academic Press.

- Choudhury, M. D., Sundara, H., John, A., Seligmann, D. D., & Kelliher, A. (2010). "*Birds of a feather*": Does user homophily impact information diffusion in social media? Available at <http://arxiv.org/abs/1006.1702>.
- Clouston, S. A. P., Verdery, A., Amin, S., & Gauthier, G. R. (2009). The structure of undergraduate association networks: A quantitative ethnography. *Connections*, 29(2), 18-31.
- Craig, E., & Wright, K. B. (2012). Computer-mediated relational development and maintenance on Facebook ®. *Communication Research Reports*, 29(2), 119-129.
- Dindia, K., & Canary, D. (1993). Definitions and theoretical perspectives on maintaining relationships. *Journal of Social and Personal Relationships*, 10(2), 163-173.
- Donath, J. S., & boyd, d. (2004). Public displays of connection. *BT Technology Journal*, 22(4), 71-82.
- Dubrofsky, R. E. (2011). Surveillance on reality television and Facebook: From authenticity to flowing data. *Communication Theory*, 21(2), 111-129.
- Ellison, N. B., Steinfield, c., & Lampe C. (2007). The benefits of Facebook "Friends:" Social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, 12, 1143-1168.
- Ellison, N. B., Steinfield, c., & Lampe C. (2010). Connection strategies: Social capital implications of Facebook-enabled communication practices. *New Media & Society*, 13(6), 873-892.
- Fiore, A. T., & Donath, J. S. (2005). Homophily in online dating: When do you like someone like yourself? *In Proceedings of the Conference on Human Factors in Computing Systems*, 1371-1374. New York: ACM Press.

- Fischer, C. S. (1982). *To dwell among friends*. Chicago: University of Chicago Press.
- Gilbert, E., & Karahalios, K. (2009). Predicting tie strength with social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 211-220. New York: ACM Press.
- Goffman, E. (1963). *Behavior in public places*. New York: The Free Press.
- Grimmelmann, J. (2009). Saving Facebook. *Iowa Law Review*, 94, 1137-1206.
- Hallinan, M. T., & Smith S. S. (1985). The effects of classroom racial composition on students' interracial friendliness. *Social Psychology Quarterly*, 48(1), 3-16.
- Hampton, K., Goulet, L. S., Rainie, L., & Purcell, K. (2011). Social networking sites and our lives. *Pew Internet & American Life Report*, June 16, 2011. Available at: <http://www.pewinternet.org/Reports/2011/Technology-and-social-networks.aspx>
- Hargittai, E. (2007). Whose space? Differences among users and non-users of social network sites. *Journal of Computer-Mediated Communication*, 13(1), 276-297.
- Harlow, S. (2012). Social media and social movements: Facebook and an online Guatemalan justice movement that moved offline. *New Media & Society*, 14(2), 225-243.
- Huston, T. L. & Levinger, G. (1978). Interpersonal attraction and relationships. *Annual Review of Psychology*, 29, 115-156.
- Homans, G. C. (1950). *The Human Group*. New York: Harcourt, Brace.

- Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2), 405-450.
- La Fond, T., & Neville, J. (2010). Randomization tests for distinguishing social influence and homophily effects. *In Proceedings of the International World Wide Web Conference*, 601-610. New York: ACM Press.
- Lampe, C., Ellison, N., & Steinfield, C. (2006). A face(book) in the crowd: Social searching vs. social browsing. *In Proceedings of the 2006 20th Anniversary Conference on the Computer Supported Cooperative Work*, 167-170. New York: ACM Press.
- Lampe, C., Ellison, N., & Steinfield, C. (2007). A familiar Face(book): Profile elements as signals in an online social network. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 435-444. New York: ACM Press.
- Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. In Berger, M. (ed.) *Freedom and control in modern society*. New York: Van Nostrand, 18-66.
- Levinson, P. (2009). *New new media*. Boston: Allyn and Bacon.
- Marsden, P. V. (1988). Homogeneity in confiding relations. *Social Networks*, 10(1), 57-76.
- Marsden, P. V. (1990). Network data and measurement. *Annual Review of Sociology*, 16, 435-463.
- Marwick, A. E., & boyd. d. (2010). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114-133.
- McLaughlin, C., & Vitak, J. (2012). Norm evolution and violation on Facebook. *New Media & Society*, 14(2), 299-315.

- McPherson, M., Smith-Lovin, L. (1987). Homophily in voluntary associations: Status distance and the composition of face-to-face groups. *American Sociological Review*, 52(3), 370-379.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415-444.
- Monge, P. R., & Contractor, N. (2003). *Theories of communication networks*. Oxford: Oxford University Press.
- Moore, K., & McElroy, J. C. (2012). The influence of personality on Facebook usage, wall postings, and regret. *Computers in Human Behavior*, 28(1), 267-274.
- Nazir, A., Raza, S., & Chuah, C-N. (2008). Unveiling Facebook: A measurement study of social network based applications. *In Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement*, 43-56. New York: ACM Press.
- Papacharissi, Z. (2009). The virtual geographies of social networks: A comparative analysis of Facebook, LinkedIn and ASmallWorld. *New Media & Society*, 11(1-2), 199-220.
- Park, H. W. (2007). Network analysis of politician blogs in South Korea: 17th National assembly members. *Korean Journal of Journalism & Mass Communication*, 51(3), 385-406.
- Park, N., Kee, K. F., & Valenzuela, S. (2009). Being immersed in social networking environment: Facebook groups, uses and gratifications, and social outcomes. *CyberPsychology & Behavior*, 12(6), 729-733.
- Park, N., Jin, B., & Jin, S.-A. (2011). Effects of self-disclosure on relational intimacy on Facebook. *Computers in Human Behavior*, 27(5), 1974-1983.

- Rogers, E. M., & Bhowmik, D. K. (1970). Homophily-heterophily: Relational concepts for communication research. *Public Opinion Quarterly*, 34(4), 523-538.
- Ross, C., Orr, E. S., Sisic, M., Arseneault, J. M., Simmering, M. G., & Orr, R. R. (2009). Personality and motivations associated with Facebook use. *Computers in Human Behavior*, 25, 578-586.
- Schrum W., Cheek, N. H. Jr., & Hunter, S. M. (1988). Friendship in school: gender and racial homophily. *Sociology of Education*, 61, 227-239.
- Sheldon, P. (2008). The relationships between unwillingness-to-communicate and students' Facebook use. *Journal of Media Psychology*, 20(2), 67-75.
- Sheldon, K. M., Abad, N., & Hinsch, C. (2011). A two-process view of Facebook use and relatedness need-satisfaction: Disconnection drives use, and connection rewards it. *Journal of Personality and Social Psychology*, 100(4), 766-775.
- Suh, B., Lee, J., & Oh, J. (2012). A study on patterns of social relationships between Twitter users. *Korean Journal of Journalism & Mass Communication*, 65(5), 88-113.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, 34, 379-440.
- Thelwall, M. (2009). Homophily in MySpace. *Journal of the American Society for Information Science and Technology*, 60(2), 219-231.
- Tong, S. T., Van Der Heide, B., Langwell, L., & Walther, J. B. (2008). Too much of a good thing? The relationship between number of friends and interpersonal impressions on Facebook. *Journal of Computer-Mediated Communication*, 13(3), 531-549.

- Van Dijck, J. (2012). Facebook as a tool for producing sociality and connectivity. *Television & New Media*, 13(2), 160-176.
- Vangelisti, A. L., & Caughlin, J. P. (1997). Revealing family secrets: The influence of topic, function, and relationships. *Journal of Social and Personal Relationships*, 14(5), 679-705.
- Verbrugge, L. M. (1977). The structure of adult friendship choices. *Social Forces*, 56(2), 576-597.
- Walther, J. B., Van Der Heide, B., Hamel, L. M., & Shulman, H. C. (2009). Self-generated versus other-generated statements and impressions in computer-mediated communication: A test of warranting theory using Facebook. *Communication Research*, 36(2), 229-253.
- Wellman, B., & Wortley, S. (1990). Different strokes from different folks: Community ties and social support. *American Journal of Sociology*, 96(3), 558-588.
- Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A review of Facebook research in the social science. *Perspectives on Psychological Science*, 7(3), 203-220.
- Xiang, R., Neville, J., & Rogati, M. (2010). Modeling relationship strength in online social networks. In *Proceedings of the 19th International Conference on World Wide Web*, 981-990. New York: ACM Press.
- Yoon, T-I. (2006). The role of interpersonal communication network in modeling planned behavior of the public issue: Understanding drinking behavior of tap water. *Korean Journal of Journalism & Mass Communication*, 50(5), 211-242.

국문 초록

상호작용 행위를 통해 본 “온라인 호모필리”

: 페이스북 “친구” 관계의 친밀도에 대한 연구

호모필리는 유사한 속성을 지닌 사람들끼리 그렇지 않은 사람들보다 더 친밀감을 느끼고 어울리는 경향 혹은 현상을 말한다. 최근 온라인상에서의 컴퓨터 매개 커뮤니케이션이 (CMC) 활발해지고 사람들이 소셜네트워크서비스를 (SNSs) 통해 사실상 “실시간”으로 다른 사람들과 상호작용을 함에도 불구하고, 호모필리에 대한 연구들은 여전히 오프라인 상황을 중심으로 진행되었다. 다양한 소셜네트워크서비스 (SNSs) 중에서도 특히 페이스북 사용자들의 “친구” 관계는 일차적으로는 기존의 오프라인 친구 관계에 기반하면서도, 전혀 모르는 사이에서부터 매우 가까운 사이에 걸쳐, 흔히 말하는 친구관계보다 훨씬 넓은 범위의 관계를 가리킨다는 점에서 독특하다. 이러한 커뮤니티의 특성에도 불구하고, 페이스북을 포함한 온라인상에서 나타나는 호모필리 현상에 대한 연구는 오프라인상에서 발전된 전통적인 호모필리 개념을 그대로 적용함으로써 이루어져 왔다. 따라서, 본 연구는 전통적인 호모필리 개념이 “온라인 호모필리”에도 그대로 적용가능한지에 대해 페이스북 사용자들의 프로필 정보와 상호작용 로그 데이터 및 친밀도에 대한 설문을 통해 살펴본다, 이에 대한 경험적 관찰을 통해서, 사람들의 유사성에 기반한 기존의 호모필리 현상이 온라인 공간에서는 나타나지 않았지만, 사용자들 사이의 상호작용 행위로

온라인에서의 호모필리 현상이 볼 수 있음을 발견했다. 이와 같은 연구결과는, 전통적인 호모필리 개념을 기반으로 한, CMC 환경에 보다 적합한 “온라인 호포밀리” 개념에 대한 정교화 작업의 필요성을 보여준다.

키워드: 호모필리, 컴퓨터 매개 커뮤니케이션 (CMC), 친밀감,
상호작용 행위, 소셜네트워크서비스 (SNSs), 페이스북, 유사성,
사용자 프로필

학번: 2011-23156